On the negative dependence inequalities and maximal score in round-robin tournament

Yaakov Malinovsky *
Department of Mathematics and Statistics
University of Maryland, Baltimore County, Baltimore, MD 21250, USA

John W. Moon †
Department of Mathematical and Statistical Sciences
University of Alberta, Edmonton, AB T6G 2G1, Canada

April 6, 2021

Abstract

We extend Huber's (1963) inequality for the joint distribution function of negative dependent scores in the round-robin tournament. As a byproduct, this extension implies convergence in probability of the maximal score in a round-robin tournament in a more general setting.

Keywords: negative correlation, probabilistic inequalities, round-robin tournaments
MSC2020: 60E15, 05C20

1 Introduction

In a classical round-robin tournament, each of n players wins or loses against each of the other $n-1$ players (Moon, 2013). Denote by X_{ij} the score of player i after the game with player $j, j \neq i$. We assume that all $\binom{n}{2}$ pairs of scores $(X_{12}, X_{21}), \ldots, (X_{1n}, X_{n1}), \ldots,$ $(X_{n-1,n}, X_{n,n-1})$ are independent. Let $s_i = \sum_{i=1, i\neq j}^{n} X_{ij}$ be the score of player i ($i = 1, \ldots, n$) after playing with all $n-1$ opponents. We use a standard notation and denote by $s_{(1)} \leq s_{(2)} \leq \ldots \leq s_{(n)}$ the order statistics of the random variables $s_1, s_2, \ldots, s_n$; and

*email: yaakovm@umbc.edu
†email: jwmoon@ualberta.ca

we denote by $s_1^*, s_2^*, \ldots, s_n^*$ normalized scores (zero expectation and unit variance) with corresponding order statistics $s_{(1)}^*, s_{(2)}^*, \ldots, s_{(n)}^*$.

Result 1 (Huber (1963)). *If $X_{ij} \in \{0,1\}, X_{ij} + X_{ji} = 1, p_{ij} = P(X_{ij} = 1) = \frac{1}{2}$, and if $n \to \infty$, then $s_{(n)}^* - \sqrt{2\log(n-1)} \to 0$ in probability.*

Huber's key observation in proving Result 1 was the following lemma concerning negative dependent scores $s_1, \ldots, s_n$, where $X_{ij} \in \{0,1\}, X_{ij} + X_{ji} = 1$, which is given verbatim below.

Lemma 1 (Huber (1963)). *For any probability matrix (p_{ij}) and any numbers $(k_1, \ldots, k_m)$, $m \leq n$, the joint cumulative distribution function of the scores $s_1, \ldots, s_m$ satisfies*

$$P(s_1 < k_1, \ldots, s_m < k_m) \leq P(s_1 < k_1) \cdots P(s_m < k_m). \tag{1}$$

In this note, we extend Huber's lemma for two models. As a byproduct, this extension implies convergence in the probability of the normalized maximal score for those models.

Model I: Round-robin tournament with draws, where the score X_{ij} equals to 1, 1/2, or 0 accordingly as a player i wins, draws, or loses a game against another player j.

Model II: Round-robin tournament, where the score X_{ij} equals $0, 1, 2, \ldots, m$, $X_{ij} = m - X_{ji} = u$ with probability $p_u, u = 0, 1, \ldots, m$ and $p_u = 1 - p_{m-u}$.

2 Round-robin tournament with draws

In the case when draws are possible, X_{ij} the score of player i after the game with player $j, j \neq i$, obtains values 1, 1/2, or 0 accordingly as player i wins, draws, or loses the game against player j. Therefore, $X_{ij} + X_{ji} = 1, i \neq j$. For any $i \neq j$ denote $p_{ij} = P(X_{ij} = 1/2)$, $q_{ij} = P(X_{ij} = 1)$, $q_{ij}^{'} = P(X_{ij} = 0)$, where $p_{ij} + q_{ij} + q_{ij}^{'} = 1$.

Lemma 2. *For any probability matrices (p_{ij}) and (q_{ij}) and any numbers $(k_1, \ldots, k_m)$, $m \leq n$, the joint cumulative distribution function of the scores $s_1, \ldots, s_m$ satisfies*

$$F(k_1, \ldots, k_m) = P(s_1 < k_1, \ldots, s_m < k_m) \leq P(s_1 < k_1) \cdots P(s_m < k_m). \tag{2}$$

Proof. The method used in the proof is similar to the method used by Huber in proving Lemma 1. Starting with two particular scores s_1 and s_2, we put $s_1 = s_1' + X_{12}$, $s_2 = s_2' + X_{21}$. X_{12} and X_{21} are dependent, but s_1' and s_2' are independent. Replacing X_{12} and X_{21} with independent random variables without changing their marginal distributions, we obtain a new joint distribution of the scores, which we denote by F'. We abuse notation by putting $p \equiv p_{12}, q \equiv q_{12}, q^{'} \equiv q^{'}_{12}$, and obtain

$$
\begin{aligned}
&F' - F \equiv F'(k_1, \ldots, k_m) - F(k_1, \ldots, k_m) \\
&= \sum_{c_1=0,1/2,1} \sum_{c_2=0,1/2,1} P\left(s_1' < k_1 - c_1, s_2' < k_2 - c_2, \ldots, s_N < k_m\right) \\
&\cdot \left[P(X_{12} = c_1)P(X_{21} = c_2) - P(X_{12} = c_1, X_{21} = c_2)\right] \\
&= \sum_{c_1=0,1/2,1} \sum_{c_2=0,1/2,1} p(c_1, c_2)\left[P(X_{12} = c_1)P(X_{21} = c_2) - P(X_{12} = c_1, X_{21} = c_2)\right] \\
&= p(0,0)qq^{'} + p(0,1)(-q^{'}(1-q^{'})) + p(1,0)(-q(1-q)) + p(1,1)qq^{'} \\
&+ p(0,1/2)q^{'}p + p(1/2,0)qp + p(1/2,1)q^{'}p + p(1,1/2)qp + p(1/2,1/2)(-p(1-p)), \qquad (3)
\end{aligned}
$$

where $p(c_1, c_2) \equiv P\left(s_1' < k_1 - c_1, s_2' < k_2 - c_2, \ldots, s_m < k_m\right)$. Now we proceed as follows: The inequalities $0 \leq s_1' < k_1$ and $0 \leq s_2' < k_2$ determine a rectangle R in the $s_1' s_2'$ plane. We partition R into nine smaller rectangles determined by the portions of two vertical lines, $s_1' = k_1 - 1$, $s_1' = k_1 - 1/2$, and two horizontal lines, $s_2' = k_2 - 1$, $s_2' = k_2 - 1/2$, that are within R. These nine "basic" rectangles we label with A, B, C, D, E, F, G, H, and I, proceeding from left to right in each row, starting with the bottom row, then moving up to the middle row and finally to the top row (see Figure 1). We adopt the convention that the points on each horizontal boundary segment of these nine basic rectangles belong to the rectangle above the segment, if there is such a rectangle; and also that the points on each vertical boundary segment belong to the rectangle to the right of the segment, if there is such a rectangle. The points on the lines $s_1' = k_1$ and $s_2' = k_2$ are not considered as belonging to any of these rectangles.

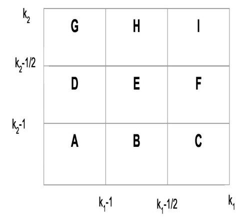

Figure 1: "Checkerboard" with three columns and three rows.

Using the above notation, we find that nine terms in (3) can be expressed as follows:

$$\begin{aligned}
&\textcircled{1} = p(0,0)qq' = qq'\,[A+B+C+D+E+F+G+H+I]\,,\\
&\textcircled{2} = p(0,1)(-q'(1-q')) = -q'(1-q')\,[A+B+C]\,,\\
&\textcircled{3} = p(1,0)(-q(1-q)) = -q(1-q)\,[A+D+G]\,,\\
&\textcircled{4} = p(1,1)qq' = qq'\,[A]\,,\\
&\textcircled{5} = p(0,1/2)q'p = q'p\,[A+B+C+D+E+F]\,,\\
&\textcircled{6} = p(1/2,0)qp = qp\,[A+B+D+E+G+H]\,,\\
&\textcircled{7} = p(1/2,1)q'p = q'p\,[A+B]\,,\\
&\textcircled{8} = p(1,1/2)qp = qp\,[A+D]\,,\\
&\textcircled{9} = p(1/2,1/2)(-p(1-p)) = -p(1-p)\,[A+B+D+E]\,.
\end{aligned}$$

Collecting the weights associated with the individual basic rectangles, we obtain

$$F' - F = \textcircled{1} + \cdots + \textcircled{9} = c_A A + c_B B + c_C C + c_D D + c_E E + c_F F + c_G G + c_H H + c_I I, \tag{4}$$

where $c_A = 0, c_B = 0, c_C = 0, c_D = 0, c_E = qq', c_F = q'(1-q'), c_G = 0, c_H = q(1-q), c_I = qq'$. It implies $F' - F \geq 0$.

The proof is completed by repeating this process for all pairs of scores. $\square$

As a byproduct, we obtain the following result:

Result 2. *If $p_{ij} = p, q_{ij} = q, q'_{ij} = q'$ and if $n \to \infty$, then $s^*_{(n)} - \sqrt{2\log(n-1)} \to 0$ in probability.*

Proof. The proof is identical to the proof in Huber (1963), replacing Lemma 1 with Lemma 2. $\square$

Remark 1. *In the symmetric case where* $q = q' = (1-p)/2$*, we obtain that* $E(s_1) = \frac{n-1}{2}, Var(s_1) = \frac{(n-1)(1-p)}{4}$*, and if* $n \to \infty$*, then*

$$s_{(n)} - \left\{ \frac{n-1}{2} + \sqrt{\frac{(n-1)(\log(n-1))(1-p)}{2}} \right\} \to 0$$

in probability.

3 Generalized round-robin tournament

In the generalized round-robin tournament, X_{ij} the score of player i after the game with player $j, j \neq i$, obtains values $0, 1, \ldots, m$ with probability $p_u = P(X_{ij} = u), u = 0, 1, \ldots, m$; $X_{ij} = m - X_{ji}$.

Lemma 3. *If* $p_u = 1 - p_{m-u}, u = 0, 1, \ldots, m$ *(symmetric distribution), then for any numbers* $(k_1, \ldots, k_m)$*,* $m \leq n$*, the joint cumulative distribution function of the scores* $s_1, \ldots, s_m$ *satisfies*

$$F(k_1, \ldots, k_m) = P(s_1 < k_1, \ldots, s_m < k_m) \leq P(s_1 < k_1) \cdots P(s_m < k_m). \tag{5}$$

Proof. Instead of considering a "checkerboard" with three columns and three rows (Figure 1), we now consider one with $m+1$ columns and $m+1$ rows. We label the columns from 0 to m from right to left, and the rows from 0 to m from top to bottom. Then we assign the coordinates uv to the square in the u-th column from the right and the v-th row from the top; the square in the top right corner has coordinates 00 and the square in the lower left corner has coordinates mm. The quantities $p(u,v)$ have essentially the same interpretation as before except that now the quantities c_1 and c_2 in (3) can range over $0, 1, \ldots, m$. From the assumption of symmetry $p_u = 1 - p_{m-u}, u = 0, 1, \ldots, m$ follows

$$p(u,v) = \begin{cases} p_u p_v & if \quad u + v \neq m \\ p_u^2 - p_u & if \quad u + v = m \end{cases} \tag{6}$$

In the expression for $F' - F$, the more complete definition of any particular term $p(u,v)$ has the effect of "exporting" its value to all squares gh such that $g \geq u$ and $h \geq v$ (i. e. to the squares farther away from the right side and the top of our checkerboard, in addition to the square uv itself). To simplify the sum of these exports from all squares uv, we instead calculate the sum of the "imports" into square gh over all squares gh. If we let $W(g,h)$ denote the sum of the probabilities imported into square gh (it corresponds to the $c_\bullet$ values in equation (4)), then it follows from the observation in the next to last sentence that

$$W(g,h) = \sum_{u \le g} \sum_{v \le h} p(u,v). \tag{7}$$

For notational convenience let $Q_k = p_0 + p_1 + \cdots + p_k$, and note that this and the fact that $p_i = p_{m-i}$ imply that

$$Q_k + Q_{m-k-1} = Q_m = 1. \tag{8}$$

- **Assertion** 1. If $g + h < m$, then $W(g,h) = Q_g Q_h \ge 0$.

  If $u \le g$ and $v \le h$, then $u + v \le g + h < m$, so the last part of (6) does not apply here. Hence, in this case it follows from (7) and the first part of (6) that

  $$W(g,h) = (p_0 + \cdots + p_g)(p_0 + \cdots + p_h) = Q_g Q_h. \tag{9}$$

  Next, if $g + h \ge m$, then we will need to take the last part of (6) into account, and we will need to "adjust" the expression $Q_g Q_h$ by subtracting the quantity p_u for every square $(u, m-u)$ in the downward-sloping diagonal of the checkerboard that coincides with or lies above and to the right of the square gh-that is, for every u such that $u \le g$ and $v = m - u \le h$. These two inequalities imply that $m - h \le u \le g$ (Note that $g - (m-h) = g + h - m \ge 0$ here, so there is certainly at least one such value of u.) Thus it follows from (6) and (7) that in this case

  $$W(g,h) = Q_g Q_h - (p_{m-h} + \cdots + p_g). \tag{10}$$

  Let us simplify this first for a special case.

- **Assertion 2.** If $g = m$ or $h = m$, then $W(m,h) = 0$ and $W(g,m) = 0$ for $0 \le h \le m$ and $0 \le g \le m$.

In this case expression (10) reduces to

$$W(m,h) = Q_m Q_h - (p_{m-h} + \cdots + p_m) = 1Q_h - (p_0 + \cdots + p_h) = Q_h - Q_h = 0.$$

And, by symmetry, $W(g,m) = 0$ also. Hence, we may assume that $g, h \leq m-1$ and, since $g + h \geq m$, that $g, h \geq 1$.

- **Assertion 3.** If $1 \leq g, h \leq m-1$ and $g + h \geq m$, then

$$W(g,h) = Q_{m-h-1} Q_{m-g-1} = (1 - Q_g)(1 - Q_h) \geq 0.$$

In this case, it follows from (9) and (10) that

$$\begin{aligned} W(g,h) &= Q_g Q_h - (p_{m-h} + \cdots + p_g) = Q_g(1 - Q_{m-h-1}) - (Q_g - Q_{m-h-1}) \\ &= Q_{m-h-1}(1 - Q_g) = (1 - Q_h)(1 - Q_g) \geq 0. \end{aligned}$$

□

Remark 2. *It is not clear if the condition of distribution symmetry is a necessary condition.*

Result 3. *If* $p_u = 1 - p_{m-u}, u = 0, 1, \ldots, m$ *and* $n \to \infty$*, then* $s^*_{(n)} - \sqrt{2\log(n-1)} \to 0$ *in probability.*

Proof. The proof is identical to the proof in Huber (1963) upon replacing Lemma 1 with Lemma 3. □

Acknowledgement

YM thanks Abram Kagan for describing a score issue in chess round-robin tournaments with draws.

References

Huber, P. J. (1963). A remark on a paper of Trawinski and David entitled: Selection of the best treatment in a paired comparison experiment. Ann.Math.Statist. **3**4, 92–94.

Moon, J. W. (2013). Topics on Tournaments. [Publicaly available on website of Project Gutenberg `https://www.gutenberg.org/ebooks/42833`].